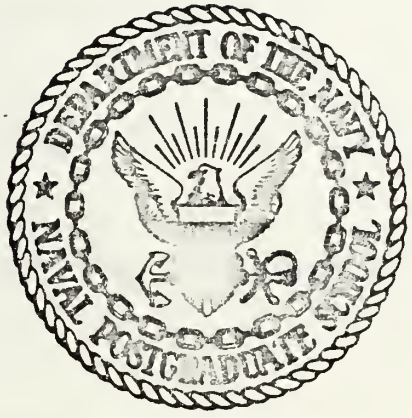
MEASUREMENT OF INCREASE IN MANAGERIAL ABILITY RESULTING FROM GRADUATE LEVEL EDUCATION OF TECHNICALLY ORIENTED FEDERAL EMPLOYEES: A REVIEW AND PROPOSAL FOR THE NAVAL AIR FEDERAL EXECUTIVE MANAGEMENT PROGRAM

Verlyne Wayne Daniels

NAVAL POSTGRADUATE SCHOOL
Monterey, California



THESIS

Measurement of Increase in Managerial Ability
Resulting from Graduate Level Education
of Technically Oriented Federal Employees:
a Review and Proposal for
the Naval Air Federal Executive Management Program

by

Verlyne Wayne Daniels

December 1975

Thesis Advisor: J. W. Creighton

Approved for public release; distribution unlimited.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Measurement of Increase in Managerial Ability Resulting from Graduate Level Education of Technically Oriented Federal Employees: a Review and Proposal for the Naval Air Federal Executive Management Program		5. TYPE OF REPORT & PERIOD COVERED Master's Thesis December 1975 6. PERFORMING ORG. REPORT NUMBER
7. Author Verlyne Wayne Daniels		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		12. REPORT DATE December 1975 13. NUMBER OF PAGES 65
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Naval Postgraduate School Monterey, California 93940		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Release; unlimited distribution

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Administrative skill	Development of managers	Management effectiveness
Assessment	Evaluation	Managerial behavior
Assessment center	Executive potential	Managerial skill
Behavior of executives	Job performance	(cont'd next page)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Background information on the Naval Air Federal Executive Management Program is provided. Managerial skills and the manager's job are discussed and described; various strategies and designs for evaluation of training and education are reviewed; a brief overview of the assessment center is given, including a description and comments on validity, on selection and training of assessors, and on design and establishment of an assessment center. A

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601 |
(Page 1)

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

19 (cont'd)

Measurement of managerial skill
Prediction of managerial skill
Selection of managers
Simulation
Situational test
Performance Dimensions

20 (cont'd)

recommendation is made that consideration be given to the use of an abbreviated assessment center for measurement of change in management skill resulting from graduate level education in the Naval Air Federal Executive Management Program (NAVAIR FEMP).

Measurement of Increase in Managerial Ability Resulting from Graduate Level Education of Technically Oriented Federal Employees: a Review and Proposal for the Naval Air Federal Executive Management Program

by

Verlyne Wayne Daniels
Captain, United States Navy
B.S., Naval Postgraduate School, 1966

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN MANAGEMENT

from the

NAVAL POSTGRADUATE SCHOOL
December 1975

ABSTRACT

Background information on the Naval Air Federal Executive Management Program is provided. Managerial skills and the manager's job are discussed and described; various strategies and designs for evaluation of training and education are reviewed; a brief overview of the assessment center is given, including a description and comments on validity, on selection and training of assessors, and on design and establishment of an assessment center. A recommendation is made that consideration be given to the use of an abbreviated assessment center for measurement of change in management skill resulting from graduate level education in the Naval Air Federal Executive Management Program (NAVAIRFEMP).

TABLE OF CONTENTS

I. OBJECTIVE

The objective of this thesis is to explore the possibility of measuring the increase in individual managerial skills or ability and the increase in potential for filling higher management positions resulting from participation in the Naval Air Federal Executive Management Program (NAVAIRFEMP).

If a thing exists, it exists in some amount. If it exists in some amount, it can be measured.

E. L. Thorndike (quoted in Isaac and Michael [29])

II. BACKGROUND

Commencing in September 1975 a thirteen month graduate-level Federal Executive Management Program (NAVAIRFEMP) was initiated at the U. S. Naval Postgraduate School. Classes for the first group were actually conducted at the Pacific Missile Test Center, Point Mugu, California, and were taught exclusively by Naval Postgraduate School (NPS) faculty through the auspices of the Office of Continuing Education. Future classes will be conducted at the U. S. Naval Postgraduate School, Monterey, California.

Appendix A (NPS Memorandum NC4 (55So)/mc of 21 August 1975) describes the program, states the objectives, and gives a course outline.

The objectives of the course were stated as follows:

> The program is designed for the federal executive who is in a middle management position. He is assumed to have experience in governmental problems from the manager's viewpoint. The NAVAIR Federal Executive Management Program is intended to review and/or teach a number of basic skills useful in general management, and to educate the student in the proper balance and integration of those skills.

This was the first graduate-level university program ever designed for civilian federal employees at the Naval Postgraduate School. The students in the program were middle managers from a number of activities in the Naval Air Systems Command (NAVAIRSYSCOM). All had undergraduate backgrounds in engineering and ranged in grade from GS-13 to GS-15.

Application for the program was open to military and civilian employees of NAVAIRSYSCOM, primarily for individuals in grades GS-13 to GS-15 for civilians, and Lieutenant Commander to Captain for military. All candidates selected for the first class were civilians.

Rating and selection of applicants was carried out by a NAVAIRSYS-COM headquarters panel comprising representatives of management, of the Naval Air Executive Institute, and of manpower management. Rating considerations included the following:

A. Identified as High Potential

B. Education: A bachelor's degree was a prerequisite, and no exceptions were made. Other considerations were number and recency of academic (credit granting) courses taken. Individuals with a high number of recent courses in business, management, or public administration were scored negatively since it was felt that this would be a duplication of effort.

C. Number of years of supervisory experience.

D. Reasons for wanting to participate: The individual's statement of intent was considered as an important factor.

E. Endorsement by Commanding Officer: The individual's endorsement by his Commanding Officer was considered critical.

F. Organizational limitation on numbers: In some cases an activity had several high-ranking candidates but could only release one or two for a program of this length. Accordingly, selection from these activities was limited to the number which could be spared even though others might have qualified.

III. MANAGERIAL SKILLS

The first step in measuring is to describe or define what one is trying to measure. This paper deals with managerial skill or ability. It is, therefore, appropriate to begin with a discussion of managers and the skills or abilities which they possess.

The literature is replete with articles and books which describe the skills, abilities, and behavior patterns which are characteristic of a successful manager or which are considered requisites for effective management. Among the more comprehensive and enlightening articles are Robert L. Katz's "Skills of an effective administrator" (first published in 1955 and reprinted with retrospective comments in 1974) and Henry Mintzberg's most recent "The manager's job, folklore and fact."

Katz [6] defines an administrator as "one who (a) directs the activities of other persons and (b) undertakes the responsibility for achieving certain objectives through these efforts." He then defines skill as "an ability to translate knowledge into action" and describes three skills which an effective administrator must possess.

Technical Skill

> ...technical skill implies an understanding of and proficiency in, a specific kind of activity, particularly one involving methods processes, procedures, or techniques....Technical skill involves specialized knowledge, analytical ability within that speciality, and facility in the use of the tools and techniques of the specific discipline.

Human Skill

> ...human skill is the executive's ability to work effectively as a group member and to build cooperative effort within the team he leads. As technical skill is primarily concerned with working with "things" (processes or physical objects) so human skill is primarily concerned with working with prople.

Conceptual Skill

> ...conceptual skill involves the ability to see the enterprise as a whole; it includes recognizing how the various functions of the organization depend on one another, and how changes in any one part affect all the others; and it extends to visualizing the relationship of the individual business to the industry, the community, and the political, social, and economic forces of the nation as a whole.

Katz noted that

> The relative importance of these three skills seems to vary with the level of administrative responsibility. At lower levels, the major need is for technical and human skills. At higher levels, the administrator's effectiveness depends largely on human and conceptual skills. At the top, conceptual skill becomes the most important of all for successful administrators.

In a retrospective commentary made when the article was reprinted several years after initial publication, Katz divides human skills into "(a) leadership ability within the manager's own unit and (b) skill in intergroup relations." Making further comments on conceptual skill and technical skill he states that conceptual skill has come to be known as the "general management point of view" and

> ...involves thinking in terms of the following: relative emphasis and priorities among competing objectives and criteria; relative tendencies and probabilities (rather than certainties); rough correlations and patterns among elements (rather than clear-cut cause-and-effect relationships)...Unless a person has learned to think this way early in life, it is unrealistic to expect a major change on reaching executive status. Job rotation, special interdepartmental assignments, and working with case problems

certainly provide opportunities for a person to enhance previously developed conceptual abilities. But I question how easily this way of thinking can be inculcated after a man passes adolescence. In this sense, then, conceptual skill should perhaps be viewed as an innate ability.

Further comments on technical skill are the following:

> In smaller companies, where technical expertise is not as pervasive and seasoned staff is not as available, I believe the chief executive has a much greater need for personal experience in the industry. He not only needs to know the right questions to ask his subordinates; he also needs enough industry background to know how to evaluate the answers.

Katz concludes that

> ...managers at all levels require some competence in each of the three skills...Dealing with the external demands on a manager's unit requires conceptual skill; the limited physical and financial resources available to him tax his technical skill; and the capabilities and demands of the persons with whom he deals make it essential that he possess human skill. A clear idea of these skills and of ways to measure a manager's competence in each category still appears to me to be a most effective tool for top management ...in the selection, training, and promotion of managers at all levels.

Mintzberg [10] explores the manager's job, exposes what he considers some myths in the historic concept, and then develops a description of the manager's job in terms of the "roles" a manager plays. These ten roles are grouped into three areas. Area one comprises INTERPERSONAL ROLES as a Figurehead, as a Leader and in the performance of Liaison. Area two comprises INFORMATIONAL ROLES as Monitor, Disseminator, and Spokesman. Area three comprises DECISIONAL ROLES as Entrepreneur, Disturbance handler, Resource allocator, and Negotiator.

These roles are further described as follows:

Figurehead:

By virtue of his position as head of an organizational unit, every manager must perform some duties of a ceremonial nature...

Duties that involve interpersonal roles may sometimes be routine, involving little serious decision making. Nevertheless, they are important to the smooth functioning of an organization and cannot be ignored by the manager.

Leader:

Because he is in charge of an organizational unit, the manager is responsible for the work of the people in that unit. His action in this regard constitutes the Leader role. Some of these actions involve leadership directly--for example, in most organizations the manager is normally responsible for hiring and training his own staff...

The influence of the manager is most clearly seen in the leader role. Formal authority vests him with great potential power; leadership determines in large part how much of it he will realize.

In the liaison role the manager makes contacts outside the vertical chain of command.

Managers spend as much time with peers and other people outside their units as they do with their own subordinates--and, surprisingly, very little with their own superiors.

...the manager cultivates such contacts largely to find information. In effect, the liaison role is devoted to building up the manager's own external information system--informal, private, verbal, but, nevertheless, effective.

Highlights of the Informational Roles are the following:

By virtue of his interpersonal contacts, both with his subordinates and with his network of contacts, the manager emerges as the nerve center of his organizational unit...

The processing of information is a key part of the manager's job. In my study, the chief executives spent 40% of their contact time on activities devoted exclusively

to the transmission of information: 70% of their incoming mail was purely informational (as opposed to requests for action). The manager does not leave meetings or hang up the telephone in order to get back to work. In large part, communication is his work. Three roles describe these informational aspects of managerial work.

As monitor, the manager perpetually scans his environment for information, interrogates his liaison contacts and his subordinates, and receives unsolicited information, much of it as a result of the network of personal contacts he has developed...a good part of the information the manager collects in his monitor role arrives in verbal form...

He must share and distribute much of this information. Information he gleans from outside personal contacts may be needed within his organization. In his disseminator role, the manager passes some of his privileged information directly to his subordinates, who would otherwise have no access to it. When his subordinates lack easy contact with one another, the manager will sometimes pass information from one to another.

In his spokesman role, the manager sends some of his information to people outside his unit--a president makes a speech to lobby for an organizational cause, or a foreman suggests a product modification to a supplier. In addition... every manager must inform and satisfy the influential people who control his organizational unit...

The manager's four roles as Decision-maker are discussed as follows:

Information is not, of course, an end in itself; it is the basic input to decision making. One thing is clear in the study of managerial work: the manager plays the major role in his unit's decision-making system. As its formal authority, only he can commit the unit to important new courses of action; and as its nerve center, only he has full and current information to make the set of decisions that determines the unit's strategy. Four roles describe the manager as decision-maker.

As entrepreneur, the manager seeks to improve his unit, to adapt it to changing conditions in the environment. In his monitor role, the president is constantly on the lookout for new ideas. When a good one appears, he initiates a development project that he may supervise himself or

delegate to an employee (perhaps with the stipulation that he must approve the final proposal)...

While the entrepreneur role describes the manager as the voluntary initiator of change, the disturbance handler role depicts the manager involuntarily responding to pressures. Here change is beyond the manager's control. He must act because the pressures of the situation are too severe to be ignored: strike looms, a major customer has gone bankrupt, or a supplier reneges on his contract...

In effect, every manager must spend a good part of his time responding to high-pressure disturbances. No organization can be so well run, so standardized, that it has considered every contingency in the uncertain environment of advance. Disturbances arise not only because poor managers ignore situations until they reach crisis proportions, but also because good managers cannot possibly anticipate all the consequences of the actions they take.

The third decisional role is that of resource allocator. To the manager falls the responsibility of deciding who will get what in his organizational unit. Perhaps the most important resource the manager allocates is his own time...

Also in his role as resource allocator, the manager authorizes the important decisions of his unit before they are implemented. By retaining this power, the manager can ensure that decisions are interrelated; all must pass through a single brain. To fragment this power is to encourage discontinuous decision making and disjointed strategy...

The final decisional role is that of negotiator. Studies of managerial work at all levels indicate that managers spend considerable time in negotiations: the president of the football team is called in to work out a contract with the holdout superstar; the corporation president leads his company's contingent to negotiate a new strike issue; the foreman argues a grievance problem to its conclusion with the shop steward...

These negotiations are duties of the manager's job; perhaps routine, they are not to be shirked. They are an integral part of his job, for only he has the authority to commit organizational resources in "real time" and only he has the nerve center information that important negotiations require.

Mintzberg states that these ten roles are not easily separable; they form an integrated whole which will not be left intact if any role is separated. For example, a manager lacking in liaison contacts as a source of external information cannot effectively make decisions that are influenced by external conditions. Neither can he disseminate the information that his subordinates need.

Lopez [14] , in his landmark work, Evaluating Executive Decision Making states:

> Managerial behavior, then, can be perceived as a pattern of responses to a steady stream of inputs presented largely in the form of letters, reports, documents to sign, telephone calls, and personal interviews. Each input requires the subject to choose from an array of alternatives available to him. The pattern of responses he typically chooses describes his style of managing and determines his effectiveness.
>
> The effectiveness of a particular response depends upon its impact on company operations, its social consequences, the amount of time available, and the relative completeness of the information necessary to make a decision. A manager's overall effectiveness can be gauged by the importance of the problems he concentrates upon; the amount of work he produces; and the appropriateness of the decisions he makes, in terms of their financial and social consequences. Putting it a bit differently, the measure of a manager is his ability to judge a situation correctly; to assess accurately the operational and social significance of a decision; to make such decisions under time pressure; and to secure the cooperation of his peers, superiors, and subordinates in the implementation of the decision.

The three skills described by Katz, the ten roles outlined by Mintzberg, and the activities discussed by Lopez are representative of the manager's job and provide criteria to be considered if one seeks to

measure improvement in managerial ability. The next step is to determine what approach to take in selecting or developing a method of measurement.

IV. MEASUREMENT AND EVALUATION

A. STRATEGIES FOR EVALUATION

Measurement of increase in skills or ability involves evaluation of the learning or change in behavior which has taken place as a result of training or education. In its "Survey of Executive Development Training Programs in Selected Number of Federal Agencies" done under contract to the Naval Aviation Executive Institute, the University of Southern California [17] outlines four levels of strategies for evaluating Executive Development Programs. These are based on Kirkpatrick's[7] four levels and are described as follows:

> Level I - The Participants Reaction Level. Basically, the question asked is: How satisfied was the participant with the program.
>
> Instruments are used to elicit the participant's reaction to the course in terms of instruction (presentation), content and relevance. Feedback may also be obtained from consultants, managers and other observers of the training effort.
>
> The instrument to measure course quality is often the "happiness" scale, i.e., rate the various criteria on a scale "1" to "10" or "poor" to "excellent" or some such basis.
>
> Course quality measurements appear to present the least amount of problems. Of course, factors such as physical surroundings, "work pile-up" back at the office induce biases into the evaluation. However, the data derived from this source is deemed by certain respondents to be good "management" data. It is an evaluation by the person who knows his or her own needs, his professional judgement, its ease in administering, and other factors which tend to make its use widespread.

Level II - The Participants Learning Level. Basically, the question asked is: What has the participant learned?

Standardized tests and self-assessment instruments are often used to determine the changes in the participant's knowledge, skill, and attitudes. The results of the tests are evaluated in terms of the learning objectives set for the course or "norms" developed from studies of other executives and managers or from job and task analyses.

Level III - The Participants Behavior Level. Basically, the question asked is: How has the participant performed differently "back home?" Various techniques and instruments are employed in Level III Evaluations:

1. Participants identify expected behavior "as a result of this training I will..." After a predetermined elapsed time, the participant will report on his progress and achievement.

2. Periodic assessment, including self-assessments, prior to training, results from assessment center activities (used by most of the agencies), and assessments subsequent to the training at various intervals--week, six months, year.

3. Observations and performance ratings over and beyond the regular employee performance rating by superiors, coaches, counselors, against predetermined standards such as ability to delegate, etc.

Level IV - The Organizational Performance Level. Basically, the question asked is: What impact does training have on the productivity or output of the organization?

Evaluation at this level involves assessing the impact of training time on the performance of the organization measured by objectively determined productivity or output measures of the organization...At this level of evaluation not only are the level III factors present but there is an added factor of identifying productivity measures, that is, answering the question: What are the outputs of the organizational unit and how are these measured?

Kirkpatrick [7], on whose work the foregoing four levels of strategy are based, defines evaluation as being "to determine the

effectiveness of a training program" and then enumerates four steps in evaluation:

Step 1--Reaction

Step 2--Learning

Step 3--Behavior

Step 4--Results

Before giving guidelines and suggested procedures for evaluating reaction he states:

> Reaction may best be defined as how well the trainees liked a particular training program. Evaluating in terms of reaction is the same as measuring the feelings of the conferees. It is important to emphasize that it does not include a measurement of any learning that takes place.

Reaction not only gives no measure of learning, it gives no indication of the change in behavior resulting from a training program.

Commenting on evaluation by measurement of learning Kirkpatrick defines learning for the purposes of his discussion as:

> the principles, facts, and techniques which were understood and absorbed...it does not include on-the-job use of these principles, facts, and techniques.

Measurement of learning thus gives no indication of changes in behavior or on-the-job application of skills learned in a training program.

In discussing evaluation of training in terms of behavior Kirkpatrick states:

> Evaluation of training programs in terms of on-the-job behavior is more difficult than the reaction and learning evaluations...A more scientific approach is needed, and many factors must be considered...

Several guideposts are to be followed in evaluating training programs in terms of behavioral changes:

1. A systematic appraisal should be made of on-the-job performance on a before-and-after basis.

2. The appraisal of performance should be made by one or more of the following groups (the more the better):
   a. The person receiving the training.
   b. His superior or superiors.
   c. His subordinates.
   d. His peers and other people thoroughly familiar with his performance.

3. A statistical analysis should be made to compare before-and-after performance and relate changes to the training program.

4. The post-training appraisal should be made three or more months after the training so that trainees have an opportunity to put into practice what they have learned. Subsequent appraisals may add to the validity of the study.

5. A control group (not receiving the training) should be used.

With regard to evaluation by measurement of results Kirkpatrick continues:

> The objectives of most training programs can be stated in terms of results such as: reduced turnover; reduced costs; improved efficiency; reduction in grievances; increase in quality and quantity of production; or improved morale...From an evaluation standpoint, it would be best to evaluate training programs directly in terms of results desired. There are, however, so many complicating factors that it is extremely difficult, if not impossible to evaluate certain kinds of programs in terms of results. Therefore, it is recommended that training directors evaluate in terms of reaction, learning and behavior.

From the foregoing it may be concluded that only Level III--The Participants Behavior Level would be of practical use in measuring a change in skill or ability resulting from training or education.

Level I--The Participants Reaction Level - is largely subjective. While it may be valuable in providing constructive criticism of a training or education program which would lead to improvements or changes in course content, it gives no indication of future on-the-job performance resulting from the training or education.

Level II--The Participants Learning Level - applies to knowledge, skills and attitudes. While it may give some indication of what a person says he or she would do in a given situation, it does not provide a good measure of what he or she actually does when confronted with a problem in a given set of circumstances.

Level IV--The Organizational Performance Level - involves measuring the output of the organization as a whole and is therefore not applicable to measurement of individual change in performance if more than one individual from an organization (or from an identifiable "performance center" within the organization) is involved. Kirkpatrick [7] states that evaluation in terms of results is extremely difficult if not impossible. The University of Southern California survey's comments on Organizational Performance Level [17] supports Kirkpatrick's position by pointing out the requirement to identify productivity measures in order to answer the question: "What are the outputs of the organizational unit and how are these measured?" Because of the variety of activities in the NAVAIRSYS-COM these measures would be extremely difficult to apply to the NAVAIR

FEMP. Of overriding importance is the fact that it is desired to measure individual improvement in managerial ability and the measurement of Results of Organizational Performance Level is not well suited to this even if it were possible. It is more applicable to overall evaluation of a program.

A further discussion of Level III follows in the next subsection.

B. EVALUATION OF TRAINING AT LEVEL III

In the previous subsection it was concluded that behavior is the only practical usable measure of change in managerial ability. As noted above, Kirkpatrick [7] states that a systematic appraisal should be made of on-the-job performance on a before and after basis and gives several guidelines concerning the appraisal technique. However, this procedure is more appropriate to situations wherein the person undergoing training returns to a previously held job and does not necessarily give a measure of ability to fill higher management positions which is the objective of this study.

Appraisal of on-the-job performance in positions of higher responsibility presents a more difficult problem since the subject has not been required to fill this position prior to training and may not be afforded the opportunity to do so for some time after the training. A suggested solution to this problem is to use a simulation of higher level positions in order to conduct an appraisal of on-the-job performance before and after training. In case it is not possible to conduct an appraisal before training, the

appraisal may be carried out after training and the performance of the trainees compared with that of a control group which has not received the training.

The use of simulation is not a unique idea as regards evaluation of performance. It has been used for some time in the Assessment Center method of executive development and/or selection for higher level positions. So far as this writer has been able to determine, it has not, however, been applied to the evaluation of individuals undergoing graduate level management education, or to the measurement of their change in managerial skill or ability.

If the assessment center is a valid method of selecting candidates for higher management positions, it must, in some way, measure the attributes or skills that are required for that higher position. If this is the case it would seem reasonable to use the assessment center technique to measure any changes which might occur as a result of training or education. This idea will be developed further in Section VI - THE ASSESSMENT CENTER.

Having determined the strategy level at which the measurement is to be carried out, the next step is to determine the design of the testing or evaluation process. The subject of design for research measurement will be discussed in the following section.

V. DESIGNS FOR RESEARCH

A. GENERAL CONSIDERATIONS

Measurement of changes in skills or ability requires some process of testing or data collection. The order and conditions under which the testing takes place or the data are collected comprise the design of the measurement process.

Campbell and Stanley [4] discuss sixteen designs for research on teaching. Of these designs, three are pre-experimental (involving pre- and post-training tests, questionnaires, interviews, and no control groups), three are true-experimental (involving control groups, statistical tests, etc.--requires random assignment to experimental group and control group), and ten are quasi-experimental (same procedures as true-experimental, but in a setting which does not allow control of all relevant factors, e.g., a study in which random assignment of comparison groups is not possible). Each design is described and then evaluated in terms of sources of errors or invalidity which might be present in the design.

These sources of invalidity are classified as to their effects on internal validity or external validity which are defined as follows:

> ...Internal validity is the basic minimum without which any experiment is uninterpretable: Did in fact the experimental treatments make a difference in this specific experimental instance? External validity asks the question of generalizability: to what populations, settings, treatment variables, and measurement variables can this effect be generalized? Both types of criteria are obviously important, even though they are frequently at odds in that features increasing one may jeopardize the other. While internal validity is the

sine qua non, and while the question of external validity, like the question of inductive inference, is never completely answerable, the selection of designs strong in both types of validity is obviously our ideal...

Relevant to internal validity, eight different classes of extraneous variables will be presented; these variables, if not controlled in the experimental design, might produce effects confounded with the effect of the experimental stimulus. They represent the effects of:

1. History, the specific events occurring between the first and second measurement in addition to the experimental variable.

2. Maturation, processes within the respondents operating as a function of the passage of time per se (not specific to the particular events), including growing older, growing more tired, and the like.

3. Testing, the effects of taking a test upon the scores of a second testing.

4. Instrumentation, in which changes in the calibration of a measuring instrument or changes in observers or scorers used may produce changes in the obtained measurements.

5. Statistical regression, operating where groups have been selected on the basis of their extreme scores.

6. Biases resulting in differential selection of respondents for the comparison groups.

7. Experimental mortality, or differential loss of respondents from the comparison groups.

8. Selection-maturation interaction, etc., which in certain of the multiple-group quasi-experimental designs, such as Design 19 (Nonequivalent Control Group Design), is confounded with, i.e., might be mistaken for, the effect of the experimental variable.

The factors jeopardizing external validity or representativeness which will be discussed are:

9. The reactive or interaction effect of testing, in which a pretest might increase or decrease the respondent's sensitivity or responsiveness to the experimental variable

and thus make the results obtained for a pretested population unrepresentative of the effects of the experimental variable for the unpretested universe from which the experimental respondents were selected.

10. The interaction effects of selection biases and the experimental variable.

11. Reactive effects of experimental arrangements, which would preclude generalization about the effect of the experimental variable upon persons being exposed to it in non-experimental settings.

12. Multiple-treatment interference, likely to occur whenever multiple treatments are applied to the same respondents, because the effects of prior treatments are not usually erasable...

Campbell and Stanley use coded graphic representations to describe these various designs. An X represents the exposure of a group to an experimental variable or event (for purposes of this paper this would be participation in a course of instruction in management); an O represents some observation or measurement (test or other procedure used to measure the effects of the instruction or education). The Xs and Os in a given row are applied to the same specific person or group. The left-to-right dimension indicates the temporal order, and Xs and Os vertical to one another are simultaneous. A symbol R is used to indicate random assignment to separate treatment groups. This randomization process occurs at a specific time and is for the purpose of achieving pretreatment equality of groups.

An example, which Campbell and Stanley call the Posttest-Only Control Group Design, is illustrated below:

R X O

R O

This represents a design in which two groups (experimental group and control group) are randomly selected, the experimental group undergoes exposure to some event, and then both groups are observed to measure the effects of the event (education) upon the experimental group.

B. PREFERRED DESIGNS

Considering the desirability of controlling as many as possible of the factors contributing to invalidity and of keeping the entire selection and testing process as simple as possible it would appear that, of the sixteen designs discussed, two of those which are classified as true experimental are feasible, suitable and preferable for use in measuring the effects of education on managerial skills. Suitability rests upon the assumption that acceptable randomization of the experimental and control groups can be achieved. (The subject of randomization of the two comparison groups will be discussed more fully below.)

The two preferred designs are the Posttest-Only Control Group Design described above and the Pretest-Posttest Control Group Design illustrated below.

R O X O

R O O

In the Pretest-Posttest Control Group Design two groups are randomly selected and observed, the experimental group undergoes exposure to some event (education), and then both groups are again observed or tested.

Both of these designs control all of the factors described as internal sources of invalidity. As far as external sources of invalidity are concerned,

multiple-treatment interference is not a relevant factor; both designs are open to question on control of the interaction effects of selection biases and the experimental variable, and on control of the reactive effects of experimental arrangements; the Posttest-Only Control Group Design controls invalidity due to reactive or interaction effects of testing while the Pretest-Posttest Control Group Design is weak in control of this factor.

Either of these two designs would be suitable and feasible for use with the FEMP if measures were undertaken to ensure randomization of comparison groups. However, the process of selection of participants in the FEMP (candidates are drawn only from high potential middle managers and the selectees for each class are the best qualified of the current applicants) precludes random selection of comparison groups and makes it necessary to resort to the use of a quasi-experimental design such as the Nonequivalent Control Group Design for the second and subsequent classes and a pre-experimental design such as the Static Group Comparison for the first class. The Nonequivalent Control Group Design and the Static Group Comparison will be discussed after discussion of randomization of comparison groups.

C. RANDOMIZATION OF COMPARISON GROUPS

In cases where an experimental group is compared to a control group for the purposes of establishing the validity of the results obtained from observing the experimental group, it is necessary, if true-experimental designs are to be used, to ensure that the two groups possess the same

characteristics or are homogeneous as regards factors which might influence, or be influenced by, the experiment. This is accomplished by randomization.

Randomization may be described as the process of selecting a group so that the characteristics of its members are not biased toward any one characteristic or set of characteristics when compared to the population from which it is drawn. If comparison groups are drawn from a population for use with a true-experimental research design, the selection should be randomized to ensure that the two groups possess the same characteristics.

Wallen and Travers [13] give a more succinct definition of randomization:

> ...Simply stated, this principle insures that if the subjects in the comparison groups are placed in the groups on a random basis all uncontrolled individual differences in variables can be assumed to balance out, provided the groups are large enough...

While randomization and true-experimental designs are not absolute necessities for meaningful measurements, they do increase validity and should be used whenever practical.

D. DESIGNS APPLICABLE TO FEMP CLASSES

Participants in the NAVAIR FEMP were drawn from applicants in accordance with the procedures described in Section II--BACKGROUND. As noted previously, the FEMP selection process precludes randomization of comparison groups. As a result, true-experimental designs could not be applied to the FEMP and recourse would have to be made to other designs for this program.

The most suitable quasi-experimental design, the Nonequivalent Control Group Design, requires a pretest and could not be used for the first class since the class has already commenced. This design could, however, be used for the second and subsequent classes if the control group was selected in time to conduct pretests on both groups (experimental and control) before the second class commences. It will, therefore, be illustrated and discussed. (The dashed line indicates a lack of randomization of comparison groups.)

```
O   X   O
----------
O       O
```

Procedures for this design are the same as for the Pretest-Posttest Control Group Design described earlier except that the comparison groups are not randomized. For FEMP application the control group could be drawn from the non-selected applicants for a given class and should include qualified applicants who were precluded from selection due to organizational limitations on numbers (see Section II--BACKGROUND).

The Nonequivalent Control Group Design is not as effective in the control of internal sources of invalidity as are the true-experimental designs. It is questionable as to control of statistical regression and weak in control of selection-maturation interaction. It is also less effective in its control of external sources of invalidity, being weak in control of reactive or interaction effect of testing and questionable in its control of interaction effects of selection biases and in its control of reactive effects of experimental arrangements. Multiple treatment

interference is not relevant. These weaknesses, although not to be ignored, do not preclude drawing useful conclusions from the results of evaluation based on this design.

None of the true-experimental or quasi-experimental designs are suitable for use with the first FEMP class. Of the pre-experimental designs, only the Static-Group Comparison illustrated below appears usable. (The dashed line indicates lack of randomization of comparison groups.)

```
X   O
------
    O
```

In this design the experimental group is exposed to some event and is then tested and compared to a non-randomized comparison group. Of the internal sources of invalidity, this design controls only history, testing, instrumentation and statistical regression. It is questionable in control of maturation and weak in control of selection, experimental mortality and selection-maturation interaction. As regards external sources of invalidity, the design is weak in control of the interaction effects of selection biases and the experimental variable. Other external sources of validity are not relevant.

Considering the relatively few factors which this design controls, it would appear that the information it would provide would be of questionable value for the purposes of the FEMP. Its only practical use would be to provide practice in testing or observation.

VI. THE ASSESSMENT CENTER

A. DESCRIPTION

The assessment center has been widely discussed and well described in the literature. Allen [18] gives a comprehensive overview of the assessment center including definition, history, and description. Further excellent descriptions and discussions are provided by Byham [1,2], Byham and Wettengel [3], Cohen and Jaffee [5], Dunnette [16], Kraut [8], and Slevin [12].

Cohen and Jaffee [5] define and describe the assessment center in the following terms:

> The assessment center is a systematic method to identify and develop managerial talent...
>
> An assessment center has a number of identifiable components. Several different types of assessment techniques may be used, including interviews, tests, peer ratings, but situational exercises are considered the most important techniques used. There are various kinds of situational exercises, but they have one characteristic in common--they simulate some critical aspects of management and bring out dimensions of performance which are important for managerial success, such as leadership style, problem-solving ability, oral communications skill, or use of delegation.

Byham and Wettengel [3] describe the assessment center as follows:

> An assessment center is a method, not a place. It involves multiple evaluation techniques, including various forms of job-related simulations, and may sometimes include interviews and psychological tests. Common job simulations include in-basked simulations, management games, group discussions, simulations of interviews with subordinates or clients, fact-finding exercises, oral-presentation exercises

and written-communications exercises. The exercises are selected to bring out behavior related to the dimensions identified by research as important to job success in the target-level positions for which the participants are being considered.

The usual assessment center lasts one to three days. Its length depends on the complexity of the jobs at the target level for which the assessment is aimed and the need for specific development insights...

Assessment center results relate to the future performance of a candidate at higher management levels, not to current job performance. By observing a participant handling the problems and challenges of the higher level jobs simulated in the exercises, assessors are able to get a feeling for how the individual would perform in a higher-level job--before the promotion...

The assessment center procedure can be thought of as a supervisory or managerial job because the exercises simulate what a supervisor or manager actually does on the job.

Byham and Wettengel describe the following typical exercises which were included in an assessment center at the state civil service level:

Assigned Role Group Discussion

In this leaderless group discussion, participants, acting as a city council of a hypothetical city, must allocate a one-million-dollar federal grant in the time allotted or make other judgements on the varying proposals offered. Each participant is assigned a point of view to sell to the other team members and is provided with a choice of projects to back and the opportunity to bargain and trade off projects for support.

Non-assigned Role Group Discussion

This exercise is a cooperative, leaderless group discussion in which four short case studies dealing with problems faced by executives working in state government agencies are presented to a group of six participants. The participants act as consultants who must make group recommendations on each of the problems. Assessors observe the

participants role in the group and the handling of the content of the discussion.

In-basket Exercise

Problems that challenge middle- and upper-level executives in state government are simulated in the in-basket exercise. These include relationships with departmental superiors, subordinates and peers, representatives of other departments, representatives of executive and legislative branches, the public, and the news media. Taking over a new job, the participant must deal with memos, letters, policies, bills, etc., found in the in-basket. After the in-basket has been completed, the participant is interviewed by an assessor concerning his/her handling of the various in-basket items.

Speech and Writing Exercises

Each participant is given a written, narrative description of a policy, event, situation, etc., and three specific situational problems related to the narrative, each requiring a written response. The participant is also required to make a formal oral presentation, based upon the background narrative description, before a simulated news conference attended by the Capitol Press Corps and interested government officials and citizens (assessors).

Analysis Problem

The analysis problem is an individual analysis exercise. The participant is given a considerable amount of data regarding a state agency's field operations, which he/she must analyze and about which he/she must make a number of management recommendations. The exercise is designed to elicit behaviors related to various dimensions of managerial effectiveness. The primary area of behavior evaluated in this exercise is the ability to sift through data and find pertinent information to reach a logical and practical conclusion.

Paper and Pencil Tests

Three different commercially-available objectively scoreable tests are included in the assessment: a reading test for self-development purposes, a reasoning-ability test, and

a personality test. The latter two are being used experimentally at present, and as with the reading test, are not made available during assessor discussions.

The foregoing examples of exercises should not be construed as limiting the scope or content of exercises to be used in an assessment center but are given as examples of what has been used in one specific application. Other applications might require different situational content depending upon the job and level to which they were to be applied. This would require a study of jobs or positions as noted in Subsection VI. E--DESIGN AND ESTABLISHMENT of an assessment center.

B. VALIDITY

If a process or technique is to be used for measuring or evaluating something, one must ask the question: is this a valid method? While the assessment center has not been proven beyond all question for every management situation, much evidence indicates that it is a valid method of evaluating managerial ability or predicting managerial success in the majority of cases. Concerning validity Byham [1] states:

> Unlike many other management development techniques that industry has widely accepted, the assessment center method has been well received partly because properly controlled research has shown it to be of value. This research has reassured both business executives and professional psychologists working in the personnel area that the assessment center method is almost certainly more valid than any other means of identifying and analyzing a candidate's management potential...
>
> ...the accumulation of research findings from a variety of types of centers lends considerable credibility to the general validity of the technique.

In a survey of 20 companies that operated centers, I uncovered some 22 studies in all that showed assessment more effective than other approaches and only one that showed it exactly as effective as some other approaches. None showed it less effective...

While the effectiveness of an assessment center has not been proven beyond the shadow of a doubt, all the research, both published and unpublished, seems to indicate that the method has more validity than other existing methods. It is in this comparison that the strength of the assessment center lies. Granted that it is not perfect, it seems that using an assessment center for identifying management potential is a sounder and fairer method than those traditionally used by management.

In a later article Byham [2] states:

Accuracy of the technique has been proven in studies conducted by AT&T, IBM, Sears Roebuck and Standard Oil (Ohio). Candidates chosen by the method have been found to be two to three times more likely to be successful at higher management levels than those promoted on the basis of supervisory judgement.

...there is strong evidence from organizations more experienced with the method that the procedure is, in general, extremely valid...

There are 22 published research studies attempting to evaluate the overall validity of assessment centers applications. Fifteen show positive results, six have such small samples as to show no results, and one study based on a very small sample indicates the assessment center is not effective. While 15 positive studies may not seem like a massive research finding, it becomes more impressive when the extremely high quality and scientific rigor of many of the studies are considered and when the research is compared with research attempting to establish results of other management selection or development programs.

Reporting on studies done by Bray and Grant [19], Wollowick and McNamara [23], and Kraut and Scott [21], Kraut [8] states:

These studies and many others lead one to conclude that assessment programs have validity in predicting those

who will move ahead in an organization. Many of the studies have flaws, but there is a constant pattern of apparent validity.

Concerning assessment centers, Moses [11] states:

> Validation studies of the process show a highly significant relationship between the assessment prediction and later management success.

The in-basket exercise is one of the most widely used exercises in assessment centers and, according to Byham [1], often the most important. He states "...where it is included, the in-basket is usually the most important exercise in an assessment center."

Meyer [9] reported on the in-basket in the following terms:

> The results of this study show that the In-Basket test might serve as a valuable aid in the selection of managers. Scores on the test proved to be related significantly to job performance ratings of managers in a concurrent validation study. Further evidence showed that the test is not just measuring managerial experience. It evidently provides a measure of managerial ability or aptitude.

While the normal assessment center consists of two or three days of exercises followed by evaluation and feed-back, an interesting experimental program carried out by AT&T was the development of a one day assessment center for the early identification of supervisory potential. This Early Identification Assessment (EIA) program was reported on by Moses [11] as follows:

> The EIA program is a process which includes both the gathering of assessment data and evaluation of supervisory potential in a single day.

Results of this program were found to correlate highly with results obtained in a more extensive assessment process, known as the Personnel

Assessment Program (PAP) widely used by the Bell System and other companies. A selected group of candidates who had been evaluated in the EIA were later evaluated in PAP.

> A group of 85 men and women were selected from a pool of 441 who had attended an EIA center...
>
> ...participants ...had not received a feedback concerning their performance in EIA.

These participants were representatives of high (32 individuals), moderate (32 individuals), and low (21 individuals) supervisory potential. A correlation of .73 was found between final ratings in the EIA program and the final ratings in the PAP for the total group. This indicates that a significant relationship exists between performance in the EIA and later performance in the PAP.

Moses [11] concluded that:

> The results of this study demonstrate that a strong relationship exists between performance in the early identification and expanded assess center process.

Since the validity of the assessment center as a predictor of management potential or ability is well established, it seems reasonable to expect that the EIA, having a high correlation with PAP, would also be a valid measure of management potential or ability. This leads to the conclusion that the difference in individual Pretest and Posttest scores on a properly designed short assessment program of the EIA type should be a valid measure of increased managerial skills and ability imparted by an educational program.

C. INCIDENTAL BENEFITS

The primary benefit resulting from use of assessment centers is, of course, more reliable selection of candidates for promotion. However, there are a number of additional benefits which have been reported. These benefits accrue to the assessors as well as the participants. Byham [2] discusses these incidental benefits as follows:

> Participation in an assessment center is a developmental experience. As can be quickly recognized, many assessment exercises such as the in-basket, management games and leaderless group discussions also are training exercises. Thus, to the extent that performance feedback is provided, participation in an assessment center is a developmental experience. In most centers...considerable performance feedback is provided during the assessment program...Participants take part in professionally led critiques of their performance in group activities, and they watch their performance in groups by means of videotape. After individually taking the in-basket for assessment purposes, they meet in small groups to share their decisions and actions with each other, to evaluate their reasoning and to broaden their repertory of responses.
>
> Even without special feedback opportunities built in, there is a great deal of evidence that most participants gain in self-insight from participating in assessment exercises and that this insight is fairly accurate. The evidence comes from comparing participant responses on self-evaluation questionnaires given after exercises with assessor evaluations. Correlations of .6 and higher based on large samples from several organizations have been found...
>
> An assessor in an assessment center benefits more than the participant in terms of direct training. Between assessor training and participation as an assessor in a center, the assessor benefits in the following ways:
>
> 1. Improvement in interviewing skills
> 2. Broadening of observation skills
> 3. Increased appreciation of group dynamics and leadership styles

4. New insights into behavior
5. Strengthening of management skills through repeated working with in-basked case problems and other simulations
6. Broadening of repertory of responses to problems
7. Establishment of normative standards by which to evaluate performance
8. Development of a more precise vocabulary with which to describe behavior.

> While many on-the-job uses can be made of these improved skills, perhaps the greatest impact is in performance-appraisal interviewing. Extensive self-report data from assessors indicate a vast improvement in both accuracy and success of appraisal interviewing.

These findings are supported by Baker and Martin [15] in their report on the Federal Executive Development Program Assessment Center. Baker and Martin state:

> One of the incidental benefits of an assessment center is an improvement in the assessor's management skills due to the assessor training and participation as an assessor in the assessment center...The assessors perceived their management skills as improved after assessor training and even more improved after participation as an assessor.

If the assessment center technique were used to measure change in managerial skill resulting from a management education program, and a design using a pretest were used, an additional beiefit to the participants should be a greater benefit from the education. This would result from greater understanding of the requirements of higher management positions, greater sensitivity, and a better understanding of one's own development needs after having participated in the pretest assessment center.

D. ASSESSOR SELECTION AND TRAINING

It is obvious that the design and content of the evaluation exercises used in an assessment center are of utmost importance to the validity and success of an assessment center. It should be no less obvious that the selection and training of assessors is also of primary importance. Allen [18] reports:

> Several sources state that familiarity with the position sought by the candidates is the single most important attribute of an assessor. This probably explains why most programs utilize line managers two or three levels above the "assessment position" as assessors. A few programs employ clinical psychologists to interpret the projective tests, but very few have psychologists actively involved in the assessment and evaluation process. Assessor/assessee ratios range from 1:1 to 1:4 and length of assignment as an assessor varies from one program (one to two weeks) to six to eight months or longer. Assessor training periods may be as short as a few days or as long as two to three weeks. The training may include lectures, films, etc., but the most effective and widely-used training technique is to involve the trainee in actual assessments on a "norm group." AT&T uses successful incumbents of the positions being sought by candidates as the "norm group." This provides assessor trainees with two benefits: (1) a tangible example of performance by organizationally defined "successes," and (2) an opportunity to err while learning without jeopardizing a candidate's career. A point upon which there is unanimous agreement is that well-trained assessors are an absolute prerequisite of an effective assessment program.

Byham [1] describes assessor selection and training as follows:

> Typically, assessors are line managers working two or three levels above the man being assessed...
>
> The job background of the assessor, of course, depends on the purpose of the specific assessment center. Where broader management aptitudes are being assessed, it is common for the assessors to be drawn from a number of areas in a company. This not only brings in a number of viewpoints, but exposes the candidates to representatives of a number

of areas where he may find promotional opportunity. Having representatives of different areas also increases the acceptance of the findings throughout the company...

A major point of controversy among operators of assessment centers is the desirability of using professional psychologists rather than specially trained managers as assessors. Most arguments for using psychologists are based on their skills in observation: they are trained to recognize behavior not obvious to the untrained eye. While this argument is plausible, it has yet to be demonstrated in an operational center. Three studies have found no differences.

However, the superiority of psychologists over completely untrained managers is well established. Becuase of this superiority, companies often use psychologists as assessors in experimental or pilot programs, where training management assessors would be difficult...

Some companies give new assessors as little as one hour of training, which really amounts to just an orientation to the whole procedure, while most others spent three or four days...

The most common method of training is by understudy. In the usual situation, an assessor-in-trining sits through an entire assessment cycle as a nonvoting member. Another method of assessor training, particularly when assessment centers are being introduced, is to have the assessors go through the assessment experience first as candidates. Everything is the same except that there are no assessors present. In a typical training situation, the assessors go through an activity such as group discussion and identify possible areas of observation afforded by the situation. Several companies videotape activities to give assessors practice in making observations.

In a separate article Byham [2] gives the following description of assessor training and assessors:

...Training for new, short-service assessors usually takes from two to five days, depending upon the complexity of the center, the importance of the assessment decision and the importance management gives to assessor training...

Assessors are usually line managers two or more levels above the participants nominated by their supervisors for the task. Line managers are used because:

1. They are familiar with the jobs for which the participant is being assessed and can therefore better judge the participant's aptitude.
2. Participation as an assessor is a developmental experience.
3. The involvement of line managers greatly increases the acceptance of the program by other managers and by the participants themselves.
4. Exposure as an assessor increases familiarity with the program, assuring most effective use of the results...

A few organizations mix line managers and personnel department or other staff people. This decision usually results from a difficulty recruiting assessors or as a means of decreasing assessor training (the trained staff people lead the line managers in completing forms, etc.).

Even less frequently are professional psychologists used as assessors...The little research available indicates that professionals do no better than trained line managers in performing their tasks. While the professional psychologists may have some superior observational skills, this is probably negated by their lack of company knowledge...

As noted above, assessors are normally selected from within an organization. In addition to the use of psychologists, there has, however, been some usage of outsiders as assessors. Byham and Wettengel [3] report:

Government jurisdictions seem to be more interested than industry in having outsiders act as assessors in their programs. The assessors may be professional psychologists or retired government executives. Aside from administrative simplification provided by this arrangement, the objectivity of outsiders seems to have particular appeal.

Outsiders may suffer from a lack of knowledge of the organization and its management and may or may not command the same respect from the participants as would internal

assessors. In some situations, outside assessors can provide more professional assessment at a cheaper cost because of the savings in training and administration expenditures. However, the use of outsiders deprives the organization's management of the substantial development benefits that come from being schooled as assessors.

It can be seen that, while outsiders may be used as assessors in some cases, it is much preferred to use managers from within the organization due to their greater familiarity with the positions being simulated and the consequent greater acceptance by the assessees. Another point in favor of using assessors from within the organization is the benefits accruing to the assessors and to the organization as discussed in the previous subsection, INCIDENTAL BENEFITS.

Having described the assessment center, discussed its validity and incidental benefits, and commented on assessor selection and training, it is appropriate to conclude the subject of assessment centers with a discussion of the design and establishment of an assessment center.

E. DESIGN AND ESTABLISHMENT

It is not within the scope of this paper to give instruction on how to design and establish an assessment center. That is a task which should be left to professional personnel managers or consultants. However, in order to provide some understanding of the magnitude of the task and of the process involved, a brief description of the procedure as outlined by Byham and Wettengel [3] is given:

The first stage in developing an assessment center is thorough job analysis to determine the major elements and to define the dimensions to be sought in the assessment center. A list of dimensions for a managerial job is not a list of the characteristics of a perfect manager; rather, the job analysis defines areas about which the assessment should be concerned--areas that should be evaluated in making a selection decision. No one is expected to be high on all of the dimensions.

In developing an assessment center...representatives of the consulting firm...and the Personnel Department staff performed the job analysis and determined the dimensions and exercises through the following steps:

1. Studied existing job descriptions of jobs in the program.
2. Studied past research in job requirements of target-level jobs and similar jobs in other organizations.
3. Interviewed a random sample of executives at the target level.
4. Conducted four meetings with groups of 4 to 8 top Civil Service and appointed executives to define critical incidents leading to success and failure in target-level jobs. It should be noted that while the executives focused on the requirements of the current target-level jobs, the executives also tried to keep an eye toward possible or probable future changes in order to develop appropriate dimensions.
5. Developed a tentative list of dimensions and definitions.
6. Administered a questionnaire to all incumbents in Career Executive positions, asking them to evaluate the importance of the dimensions to success in target-level jobs and to check the clarity of the definitions of the dimensions.
7. Developed a working list of dimensions.
8. Revised dimensions based on experience and feedback from assessor training and early pilot assessment centers.
9. Selected and developed assessment exercises to bring out the desired dimensions.
10. Trained assessors to observe and record the dimensions.

From the foregoing it can be seen that the establishment of an assessment center requires a thorough analysis of the position or position level to be simulated in order to determine what dimensions should be measured. This is followed by development of appropriate exercises and training of assessors. It is obvious that these are tasks that should be performed by professional personnel managers or consultants. Personnel from within the organization may be used if guided and supervised by professional personnel managers or consultants.

VII. FINDINGS AND CONCLUSIONS

The following are findings and conclusions reached as a result of the foregoing study of managerial skills, evaluation strategies, research designs, and assessment center processes.

1. Only Level III--The Participants Behavior Level--is of practical use in measuring a change in on-the-job demonstration of skill or ability resulting from training or education.

2. The assessment center has been validated as a method of evaluating potential on-the-job performance in a given position and of selecting candidates to fill higher management positions. This is accomplished by evaluating a number of dimensions which have been found to be representative of the managerial skills and requirements described in Section III--MANAGERIAL SKILLS.

3. While most assessment centers have been designed for application to a specific position, some have been designed for application to a general level of management within a given field. It should, therefore, be possible to design an assessment center to evaluate general managerial ability which would be applicable to the Naval Air Systems Command. Such an assessment center could be used as a vehicle to measure change in managerial ability resulting from participation in the NAVAIR FEMP using the design suggested in Section V--DESIGNS FOR RESEARCH.

4. Evaluations obtained from an abbreviated assessment center of the Early Identification Assessment type correlate highly with those obtained from more extensive assessment programs. Therefore, cost savings could be realized by using an abbreviated assessment center for the purpose of measuring change in managerial skills.

5. The difference in individual pretest and posttest scores on a properly designed short assessment program would be a valid measure of increase in managerial skills resulting from graduate level education.

6. In order to minimize bias, assessors should not review the results of pretraining assessment of participants prior to conducting posttraining assessment. Comparison of the two assessments would, of course, be required after completion of the second assessment in order to provide a measure of change in managerial abilities displayed.

7. Design of an assessment center for use with the NAVAIR FEMP would involve a job analysis of positions in the NAVAIRSYSCOM in order to develop the specific exercises to be used for evaluation.

8. For the job analysis, design of exercises, training of assessors, and initial conduct of the evaluations, it would be advisable to engage professional consultants. Sufficient "in-house" expertise should be developed from the first evaluations to allow later assessments to be conducted by NAVAIRSYSCOM/USNPS personnel and faculty.

9. Assessors should be drawn from higher management levels within the NAVAIRSYSCOM or from higher level retired NAVAIRSYSCOM executives. Insofar as possible, the same assessors should be used

for the first few times assessments are conducted in order to assure constancy of evaluation standards and provide the most accurate measurement of changes in managerial skill.

VIII. RECOMMENDATIONS

Based on the foregoing study and the resulting findings and conclusions it is recommended that:

1. Consideration be given to the development and use of an abbreviated (one day) assessment center as an instrument to measure change in management skill resulting from participation in the NAVAL AIR FEDERAL EXECUTIVE MANAGEMENT PROGRAM;

2. The assessment center be developed for application to a general level of management higher than the highest level held by the participants in the NAVAL AIR FEDERAL EXECUTIVE MANAGEMENT PROGRAM;

3. Insofar as personnel resources permit, assessors be drawn from higher level management within the Naval Air Systems Command and that other assessors be drawn from higher level retired NAVAIRSYSCOM executives;

4. For the reasons cited in Subsection V. D--DESIGNS APPLICABLE TO FEMP CLASSES, no effort be made to conduct measurement of the first class; and

5. That the Nonequivalent Control Group Design be used to measure the change in managerial skills of the second and subsequent classes in the FEMP.

APPENDIX A

NAVAL POSTGRADUATE SCHOOL
Monterey, California

NC4(55So)/mc
21 August 1975

MEMORANDUM

From: D. A. Schrady, Chairman, OR/AS Dept.
To: Academic Council

Subj: Federal Executive Management Program

Encl: (1) Program Description

1. Background. In September 1974, principals in NAVAIR civilian manpower development and CAPT Schwartz, XO of the Pacific Missile Test Center, met with the Superintendent and other NPS personnel to discuss the possibility of an NPS Masters level program in management for NAVAIR middle managers. The program description enclosed has been developed over the past year and in consideration of NAVAIR objectives and constraints, and NPS capabilities and standards. A Department curriculum committee consisting of Professors Barr, Creighton, Darbyshire, Dean, Giauque and Jolly was responsible for evolving the program structure.

2. Structure. The program as structured follows the present Administrative Science curricula and would lead to the M.S. in Management degree. The program will have a full-time director who will function in this capacity for the duration of the thirteen-month program. The director, Professor J. W. Creighton, has been involved with the program from the earliest discussions, worked with NAVAIR on student selection, and has interviewed all of the prospective students. Twenty-eight candidates were formally submitted to the Dean of Curricula for admission. The students are NAVAIR middle managers from a number of activities. They range from GS-13 to GS-15 in grade, and their undergraduate backgrounds are in engineering. All students have done some graduate work or continuing education in management subjects. Further, all students are persons identified by NAVAIR as high-potential performers.

3. Facilities. For a variety of reasons the 1975-76 program will be presented at the Pacific Missile Test Center, Pt. Mugu. The program will be taught by NPS faculty exclusively, through the auspices of Continuing Education. Faculty salaries, travel and per diem are paid by the NAVAIR Executive Institute, headquartered at Pt. Mugu. PMTC library facilities, especially in the management area, have been reviewed and judged adequate to good. Classrooms, offices and other facility support items have been provided for the program by PMTC.

NC4(55So)/mc
21 August 1975

Subj: Federal Executive Management Program

4. The program, in our opinion, satisfies the requirements for the existing M.S. in Management degree and is exciting and attractive to both NAVAIR and the Naval Postgraduate School. Academic Council approval of the proposed degree program is hereby requested. The course, MN 3002, will be submitted to the Course Advisory Committee shortly.

D. A. SCHRADY

Exhibit Number

FY 76

NAVAL POSTGRADUATE SCHOOL
Monterey, California

NAVAIR FEDERAL EXECUTIVE MANAGEMENT CURRICULUM

Background

The contemporary Navy civilian manager must have technological competence, but he must also master the science of management, embracing the organization, staffing, directing, planning and financial management of his enterprise. Skills and knowledge in contracting, production planning and scheduling, operations research and government political and budgetary processes are also essential, particularly for those involved in weapon systems acquisition. It has been amply demonstrated in the civilian economy that technologically based managers require managerial education. The magnitude of the knowledge to be mastered, plus the evidence that management can be taught most effectively to mature persons after they have had some practical leadership experience, dictates a graduate level of instruction. Graduate education in management significantly improves the performance of the manager.

A new four to six quarter Administrative Science program (the length depending on the student's background) leading to the degree of Master of Science in Management was developed at NPS and implemented in January of 1973. The NAVAIR Federal Executive Management Program satisfies the degree requirements of these existing management curricula and also responds to the needs and interests specified by the Naval Air Systems Command.

General Description

The Executive Management Program requires one year of graduate work at the Naval Postgraduate School or Postgraduate School Extension. Entrance must be preceded by completion or validation of the equivalent of a half year of work in prerequisite subject areas. These prerequisite areas represent 26 credit hours of management fundamentals which are described in further detail below. In addition, 56 credit hours of upper division graduate courses in residence are required, including at least twelve (12) quarter hours at the 4000 level, for a total of 82 credit hours.

Objective

The program is designed for the federal executive who is in a middle management position. He is assumed to have experience in governmental problems from the manager's viewpoint. The NAVAIR Federal Executive Management Program is intended to review and/or teach a number of basic skills useful in general management, and to educate the student in the proper balance and integration of those skills. The specific areas in which basic knowledge and skills will be developed are as follows:

I. Human behavior and organization theory

This includes an understanding of human behavior in organizations, methods of motivating and controlling people, organizational design issues and theory, and management/ labor relations.

II. Applied economic concepts

Required in this area is an understanding of basic economic concepts, theories, and terminology, plus a familiarity with the economic structure of the United States defense industry and the United States government.

III. Financial reporting, control, and instruments

The graduate should understand basic accounting and know how to interpret common financial reports. He should be familiar with the basics of financial control systems and financial management, industrial cash management issues and techniques, capital budgeting, the government budgeting cycle, and the major types of contracts currently used in government.

IV. Analytical decision-making tools and management information systems

This area includes management information systems, decision analysis and basic concepts in statistics and statistical analysis. Also required are such major topics in operations research as: linear, nonlinear and dynamic programming; simulation; PERT and CPM methods; queueing theory. Procurement topics such as CSCS/C will also be discussed.

V. Policy formulation and long-range planning

This area integrates the skills acquired in the functional topics and stresses planning and strategy formulation in response to external pressures and developments. Futuristic issues related to minority affairs, ecology, energy problems, population growth, world power shifts, detente, etc. will be examined.

Degree Requirements

a. Fifty-six (56) credit hours of approved graduate level work taken in residence at the Naval Postgraduate School or through an extension facility of the school, including at least 12 hours of 4000 level courses.

b. Completion or validation before entrance into the program of twenty-six (26) hours of prerequisite course work in Management Fundamentals as follows:

Human behavior and organization theory - 8 hours

A course in basic psychology or sociology or its equivalent; and a course in organization theory and practice. These are the equivalent of:

MN 2106 - Individual and Group Behavior — 4 hrs
Prerequisite: None

MN 3105 - Organization and Management — 4 hrs
Prerequisite: MN 2106

Applied economic concepts - 4 hours

Understanding of basic micro and macro economics, the equivalent of:

MN 2031 - Economic Decision Making — 4 hrs
Prerequisite: MA 2305 concurrently

Financial reporting and control - 4 hours

Ability to understand and construct profit and loss statements and balance sheets and know the basics of cash flow forecasting and management, the equivalent of:

MN 2150 - Financial Accounting — 4 hrs
Prerequisite: None

Analytical and decision-making tools - 10 hours

Understanding of differential and integral calculus, and such statistical concepts as uncertainty, mean, standard deviation, expected values, and the major properties of the most important types of statistical distributions (normal, uniform, bimodal, Poisson), the equivalent of:

MA 2305 - Differential Calculus — 3 hrs
Prerequisite: College algebra

MA 2306 - Integral Calculus — 2 hrs
Prerequisite: MA 2305 or equivalent

MA 2040 - Matrix Algebra — 2 hrs
Prerequisite: College algebra

PS 3005 - Probability — 3 hrs
Prerequisite: MA 2305 or equivalent

The above 26 hours of management fundamentals can be satisfied by any of the following methods:

1) completion of a credit course in the area given by a recognized and qualified institution;

2) passing a preliminary examination in the area; or

3) enrolling in and passing the appropriate one-month course given by the Naval Postgraduate School. The intensive one-month Continuing Education courses will be offered during a preparatory period immediately prior to the opening of the program, during which time an enrollee would be able to complete up to two of the prerequisite requirements. These are Continuing Education courses already in the catalogue or planned which will satisfy prerequisites needed for later courses.

Responsibility for determining whether or not an enrollee has satisfied the prerequisite 26 hours of management fundamentals rests upon the Naval Postgraduate School.

c. The completion of a thesis on a topic or issue typically confronting managers at executive levels.

The residence requirements will be met by completion of the following program:

Quarter I		Credit Hours
MN 3161	Managerial Accounting Prerequisite: MN 2150	4
MN 3211	Operations Analysis for Management Prerequisites: MA 2306 and PS 3005	4
MN 3002	Research Methodology (Adaptation from MN 3001; will have PS 3005 and MN 3211 concurrently as prerequisites.)	4
MN 3125	Organizational Behavior and Administration Prerequisite: MN 3105	4
	Total	16

Quarter II		
MN 3140	Microeconomic Theory Prerequisites: MN 2031, MA 2305 and MA 2040 or their equivalents	4
MN 3183	Management Information Systems and the Computer Prerequisites: PS 3005, MN 3105, MN 3211 and CS 0113	4
MN 3127	Selected Topics in Organization and Management Prerequisite: Departmental Approval	4
ELECTIVE	(see below for examples)	4
	Total	16

Quarter III		Credit Hours
MN 3101	Personnel Management and Labor Relations Prerequisites: MN 3105, MN 3140, MN 3161	4
MN 3172	Public Policy Processes Prerequisites: MN 3105, MN 3140, MN 3161	4
MN 4165	Selected Topics in Accounting and Financial Management Prerequisite: A background of advanced work in accounting and financial management and Departmental approval.	4
THESIS		-
	Total	12

Quarter IV		
MN 3371	Procurement/Contract Administration Prerequisites: MN 3140 or MN 3141	4
MN 4105	Management Policy Prerequisite: Open only to students in their final quarter of the Management Program.	4
ELECTIVE	(Must be 4000 level.)	4
THESIS		-
	Total	12
	Total Credit Hours in Residence	56

Suggested Electives

Personnel Management and Organizational Theory

MN 3110	Individual Behavior
MN 3111	Industrial Psychology
MN 3120	Planning and Control
MN 3121	Leadership and Group Behavior
MN 3124	Analysis of Bureaucracy
MN 4112	Personnel Selection and Classification
MN 4113	Personnel Training and Development
MN 4114	Personnel Performance Evaluation
MN 4115	Personnel Motivation
MN 4121	Organization Theory
MN 4123	Organization Development
MN 4147	Industrial Relations

Procurement, Material/Logistics, and Production Management

MN 3373 Transportation Management
MN 4373 Transportation Policy
MN 4371 Military Procurement Policy
MN 4172 Marketing Strategy
MN 3372 Physical Distribution and Supply Systems
MN 4376 Seminar in Material/Logistics
MN 3374 Management of Defense Production

Financial Management

MN 3251 Accounting Theory and Standards
MN 4153 Seminar in Accounting and Control
MN 4162 Cost Accounting
MN 4151 Internal Control and Auditing
MN 4161 Controllership
MN 4152 Decision Making for Financial Management
MN 4154 Seminar in Financial Management
MN 4181 Applications of Management Information Systems

MN 3002 Research Methodology

The program provides a course in Research Methodology (MN 3002). During the six month pre-campus preparation period, the student will be encouraged to locate several potential thesis subjects related to his aspirations. The course will stress:

1) Subject Selection
2) Research Planning
3) Research Methodologies
4) Presentation of Findings

The prime objective of this course will be to help the student get a firm start on his or her thesis and establish a realistic schedule in order to complete it. As a part of the thesis work, all students will be given the opportunity to spend one week with an executive in a private sector organization or in a large-scale agency. This experience visit should assist the student with his or her thesis planning and preparation, and should provide contact with a potential user of the results of the research work. At the end of the course each student is expected to have a thesis topic selected and approved, a thesis advisor located, and an outline of the proposed thesis, together with a timetable, submitted. Course documentation for MN 3002 will be developed and submitted to the Course Advisory Committee.

Thesis Preparation Time

The course in Research Methodology and the early selection of a thesis topic, together with the support provided to the student by the full-time NAVAIR Program Director, make it possible to reduce the number of scheduled thesis time slots from three to two, one each during the last two quarters of residence.

LIST OF REFERENCES

Articles

1. Byham, William C., "Assessment centers for spotting future managers," Harvard Business Review, p. 150-160, July-August 1970.

2. Byham, William C., "The assessment center as an aid in management development," Training and Development Journal, p. 10-21, December 1971.

3. Byham, William C., and Weggengel, Carl, "Assessment centers for supervisors and managers: an introduction and overview," Public Personnel Management, p. 352-364, September-October 1974.

4. Campbell, Donald T., and Stanley, Julian C., "Experimental and quasi-experimental designs for research on teaching," in Nathaniel L. Gage (ed.), Handbook of Research on Teaching, 1. 171-246, Rand McNalley & Company, 1963.

5. Cohen, Barry M., and Jaffee, Cabot L., "Assessment centers aid government agencies," Government Executive, p. 42-44, October 1973.

6. Katz, Robert L., "Skills of an effective administrator," Harvard Business Review, p. 90-102, September-October 1974.

7. Kirkpatrick, Donald L., "Evaluation of training," in Robert L. Craig and Lester R. Bittel (eds.), Training and Development Handbook, p. 87-112, McGraw-Hill Book Company, 1967.

8. Kraut, Allen I., "A hard look at assessment centers and their future," Personnel Journal, p. 317-326, May 1972.

9. Meyer, Herbert H., "The validity of the in-basket test as a measure of managerial performance," Personnel Psychology, Vol. 23(3), p. 297-307, Autumn 1970.

10. Mintzberg, Henry, "The manager's job; folklore and fact," Harvard Business Review, p. 49-61, July-August 1975.

11. Moses, Joseph L., "The development of an assessment center for the early identification of supervisory potential," Personnel Psychology, Vol. 26(4), p. 569-580, Winter 1973.

12. Slevin, Dennis P., "The assessment center: breakthrough in management appraisal and development," Personnel Journal, p. 255-261, April 1972.

13. Wallen, Norman E., and Travers, Robert M. W., "Analysis and investigation of teaching methods," in Nathaniel L. Gage (ed.), Handbook of Research on Teaching, p. 448-505, Rand McNalley & Company, 1963.

Books

14. Lopez, Felix M., Jr., Evaluating Executive Decision Making: The In-basket Technique, American Management Association, Inc., 1966.

Reports

15. United States Civil Service Commission, Technical Memorandum 74-4, Evaluation of the Federal Executive Development Program Assessment Center, by Dale R. Baker and Charles G. Martin, September 1974.

16. Office of Naval Research, Technical Report Number 4000, Multiple Assessment Procedures in Identifying and Developing Managerial Talent, by Marvin D. Dunnette, August 1970.

17. University of Southern California Washington Public Affairs Center, Survey of Executive Development Training Programs in Selected Number of Federal Agencies, by Delroy L. Cornick and James F. Wolf, Undated.

Theses

18. Allen, Galen B., The Personnel Assessment Center: A Review and Suggested Application, Masters Thesis, Naval Postgraduate School, Monterey, California, June 1974.

BIBLIOGRAPHY

Articles

19. Bray, Douglas W., and Grant, Donald L., "The assessment center in the measurement of potential for business management," Psychological Monographs: General and Applied, Vol. 80 (17), 1966.

20. Business Week, 28 August 1971, p. 34-35, "When they make believe they're boss."

21. Kraut, Allen I., and Scott, G., "The validity of an operational management assessment program," Journal of Applied Psychology, Vol. 56(2), p. 124-129, April 1972.

22. Oncken, William, Jr., and Wass, Donald L., "Management time: Who's got the monkey," Harvard Business Review, p. 75, November-December 1974.

23. Wollowick, H. B., and McNamara, W. J., "Relationship of the components of an assessment center to management success," Journal of Applied Psychology, Vol. 53, p. 348-352, 1969.

Books

24. Anderson, Scarvia B., and others, Anatomy of Evaluation: Important Concepts and Techniques in Evaluating Education/Training Programs, Educational Testing Service, Princeton, N. J., 1973.

25. Byers, Kenneth T., (ed.), Employee Training and Development in the Public Service, Public Personnel Association, 1970.

26. Craig, Robert L., and Bittel, Lester B., (eds.), Training and Development Handbook, McGraw-Hill Book Company, 1967.

27. Folsom, Marion B., Executive Decision Making, McGraw-Hill Book Company, 1962.

28. Gage, Nathaniel L., (ed.), Handbook of Research on Teaching, Rand McNalley & Company, 1963.

29. Isaac, Stephen, and Michael, William B., Handbook in Research and Evaluation, Robert R. Knapp, Publisher, 1971.

30. Jaffee, Cabot L., Problems in Supervision: An In-basket Training Exercise, Addison-Wesley Publishing Company, 1968.

31. Jaffee, Cabot L., Effective Management Selection: The Analysis of Behavior by Simulation Techniques, Addison-Wesley Publishing Company, 1971.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Documentation Center Cameron Station Alexandria, Virginia 22314	2
2.	Library, Code 0212 Naval Postgraduate School Monterey, California 93940	2
3.	Department Chairman, Code 55 Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	2
4.	Professor J. W. Creighton, Code 55Cf Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	10
5.	CDR G. B. Allen, USN, Code 55Al Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	1
6.	CAPT Verlyne W. Daniels, Code 21D Naval Facilities Engineering Command 200 Stovall Street Alexandria, Virginia 22332	1
7.	Mr. John H. Dent Department of the Navy Naval Air Systems Command Naval Aviation Executive Institute Washington, D. C. 20361	1
8.	Ms. Bonita M. Moore, Code 90A1 Department of the Navy Naval Air Systems Command Washington, D. C. 20361	1

9. Professor R. S. Elster 1
1667 West Wind Way
McLean, Virginia 22101